

FUNDAÇÃO UNIVERSIDADE FEDERAL DE RONDÔNIA
DEPARTAMENTO DE MATEMÁTICA E ESTATÍSTICA
BACHARELADO EM ESTATÍSTICA

KARLA SALES BARBOSA

APLICAÇÃO DA METODOLOGIA BOX & JENKINS PARA
PREVISÃO DO LEITE CRU DO ESTADO DE RONDÔNIA

JI-PARANÁ

2016

KARLA SALES BARBOSA

**APLICAÇÃO DA METODOLOGIA BOX & JENKINS PARA
PREVISÃO DO LEITE CRU DO ESTADO DE RONDÔNIA**

Trabalho de Conclusão de Curso submetido ao Departamento de Matemática e Estatística, da Fundação Universidade Federal de Rondônia, *Campus* de Ji-Paraná, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientadora: Prof.^a Roziane Sobreira dos Santos

JI-PARANÁ

2016

Dados Internacionais de Catalogação na Publicação
Biblioteca Setorial - UNIR/Campus de Ji-Paraná

B238a
2016

Barbosa, Karla Sales

Aplicação da metodologia Box & Jenkins para previsão do leite
cru do estado de Rondônia / Karla Sales Barbosa; orientador Roziane
Sobreira dos Santos. -- Ji-Paraná, 2016
33, p. : 30 cm

Trabalho de conclusão de Curso Bacharel em Estatística. –
Universidade Federal de Rondônia, 2015

Inclui referências

I. Produção de leite. 2. Pecuária Leiteira. 3. Método Estatístico.
I. Santos, Roziane Sobreira dos. II. Universidade Federal de
Rondônia. III. Título

CDU 519.2:636.2.034



ATA DA SESSÃO DE APRESENTAÇÃO E DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO

Aos 30 (trinta) dias do mês de junho de 2016, realizou-se na Sala Laboratório de Estatística 2, no Campus de Ji-Paraná, a Sessão de Apresentação e Defesa do Trabalho de Conclusão de Curso (TCC), com o título “**Aplicação da Metodologia Box & Jenkins para Previsão do Leite Cru do Estado de Rondônia**”, apresentado pela acadêmica **Karla Sales Barbosa**. Os trabalhos foram instalados às 15h02 pela presidente da Banca Examinadora, aprovada pelo Departamento e constituída por: Prof^o **Dr. Dilson Henrique Ramos Evangelista** (Universidade Federal de Rondônia - Departamento de Matemática e Estatística), Prof^o **Gr. Bruno Soares de Castro** (Universidade Federal de Rondônia - Departamento de Matemática e Estatística) e Prof^a **Dr^a Roziane Sobreira dos Santos** (Universidade Federal de Rondônia - Departamento de Matemática e Estatística), sendo esta a orientadora da acadêmica. A Banca Examinadora, tendo decidido aceitar o Trabalho de Conclusão de Curso, após a apresentação passou à arguição pública da acadêmica. Encerrando os trabalhos de arguição às 16:00 horas. A Banca Examinadora deu parecer final **Aprovada**, com a nota 90 (**noventa**), resultado da média aritmética das notas individuais atribuídas pelos membros da Banca Examinadora. Proclamado o resultado final pela presidente da Banca, foram encerrados os trabalhos. Para constar, lavrou-se a presente Ata, que é assinada pelos membros da Banca Examinadora e pela acadêmica.

Ji-Paraná, 30 de junho de 2016.

Dr. Dilson H. Ramos Evangelista
Prof. Membro da Banca

Gr. Bruno Soares de Castro
Prof. Membro da Banca

Dr^a. Roziane Sobreira dos Santos
Prof^a Orientadora e Pres. Banca

Karla Sales Barbosa
Acadêmica

A Deus

A minha mãe

Ao meu querido filho

DEDICO

AGRADECIMENTOS

A Deus.

A toda a minha família, em especial a minha mãe Maria Sales dos Santos e ao meu querido filho Heitor Sales de Oliveira, pela paciência, amor e carinho.

A todos os meus professores do Curso de Bacharelado em Estatística da UNIR.

Quero agradecer também em especial, a minha orientadora, professora Dr.^a Roziane Sobreira dos Santos, pela compreensão, paciência e principalmente pela sua amizade.

Aos colegas da graduação em Estatística pelas boas conversas, motivações e companheirismo, em especial aos mais próximos que de uma maneira, contribuíram para alcançar a conclusão do curso.

Ao Departamento de Matemática e Estatística da Universidade Federal de Rondônia – Ji-Paraná.

A Universidade Federal de Rondônia por ter me proporcionado a oportunidade de ter alcançando a primeira graduação de nível superior.

Meus sinceros agradecimentos a todos aqueles que de alguma maneira contribuíram para que a conclusão deste trabalho tornasse possível.

“Não estamos preparados para lidar com o aleatório e, por isso, não percebemos o quanto o acaso interfere em nossas vidas.”

Leonard Mlodinow, professor e escritor.

RESUMO

Ao longo dos anos, a forma do produtor explorar a produção dos bovinos foi modificando-se. A ordenhadas nas vacas é um exemplo desse processo, com uma sucessão de equipamentos cada vez mais potentes. No Brasil o agronegócio do Leite e seus derivados exercem um papel de grande prestígio no fornecimento de alimentos e na criação de empregos e renda para a população. A utilização da estatística como forma de tomada de decisão é de grande importância, gerando resultados que possam auxiliar os governantes e produtores rurais nas tomadas de decisões e realizar planejamentos estratégicos futuros. Para tais estudos foi utilizada análise de séries temporais para obter um modelo estatístico que estime a produção de leite cru no estado de Rondônia. O modelo SARIMA (1,0,0) x (1,1,0)₁₂ obteve os melhores resultados, apresentado o menor erro de acordo com os indicadores selecionados, além de resíduos com características de ruído branco, o que corrobora a qualidade do ajuste.

Palavras-chave: produção; leite; séries; SARIMA.

ABSTRACT

APPLICATION OF THE METHODOLOGY BOX & JENKINS FOR PREDICTION OF RONDÔNIA STATE RAW MILK.

Over the years, the shape of the producer explore the production of the bovines was modifying itself. The cows milked in this process is an example with a succession of increasingly more powerful equipment. In Brazil, agribusiness Milk and milk products play a prestigious role in providing food and creating jobs and income for the population. The use of statistics as a means of decision-making is of great importance, generating results that can help governments and farmers in decision-making and make future strategic planning. For such studies was used time series analysis for a statistical model that estimates the production of raw milk in the state of Rondônia. The SARIMA model $(1,0,0) \times (1,1,0)_{12}$ achieved the best results, presented the lowest error according to the selected indicators, as well as waste with white noise characteristics, which confirms the quality of adjustment.

Keywords: production; milk; series; SARIMA.

LISTA DE ILUSTRAÇÕES

Figura 1. Ciclo iterativo de Box & Jenkins	16
Figura 2. Representação gráfica da série original da produção de leite cru do estado de Rondônia, em litros, no período de janeiro 2009 até dezembro de 2014, com os respectivos meses identificados.....	22
Figura 3. Representação gráfica das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) da série original da produção de leite cru do estado de Rondônia, em litros, no período de janeiro 2009 até dezembro de 2014.....	23
Figura 4. Representação gráfica das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) da série diferenciada sazonalmente da produção de leite cru de Rondônia, em litros, no período de janeiro 2009 até dezembro de 2014.	24
Figura 5. Comparação entre os valores reais e previstos da série da produção de leite cru para o estado de Rondônia para o período de janeiro de 2015 até dezembro de 2015.....	25

LISTA DE TABELAS

Tabela 1. Resultados dos <i>p</i> -valor para os testes dos resíduos.	24
Tabela 2. Estatísticas dos modelos SARIMA (p,d,q) x (P,D,Q)s.....	25
Tabela 3. Valores reais e valores previstos do modelo SARIMA (1,0,0) x (1,1,0) ₁₂ , no período de janeiro a dezembro de 2015, da produção mensal de leite de Rondônia (em litros).	26

LISTA DE ABREVIATURAS

AIC	Critério de Informação de Akaike
ACF	Função de Autocorrelação
ADF	Dickey e Fuller Aumentado
MAPE	Erro Percentual Absoluto Médio
MQO	Mínimos Quadrados Ordinários
PACF	Função de Autocorrelação Parcial
SARIMA	Autorregressivo Integrado de Médias Móveis Sazonal

SUMÁRIO

1. INTRODUÇÃO	14
2. REFERENCIAL TEÓRICO	15
2.1 AGRONEGÓCIOS LEITEIROS	15
2.2 SÉRIES TEMPORAIS	15
3. MATERIAIS E MÉTODOS	18
3.1 BASE DE ESTUDO	18
3.2 METODOLOGIA BOX & JENKINS	19
3.3 CRITÉRIOS PARA SELEÇÃO E DESEMPENHO DOS MODELOS	21
3.4 SOFTWARE R	21
4. RESULTADOS E DISCUSSÕES	22
5. CONCLUSÕES	27
REFERÊNCIAS	28
APÊNDICE A - ROTINA DA ANÁLISE DE SÉRIES TEMPORAIS	30

1. INTRODUÇÃO

Ao longo dos anos, a forma do produtor explorar a produção dos bovinos foi modificando-se. As ordenhadas nas vacas é um exemplo desse processo, com uma sucessão de equipamentos cada vez mais potentes (MAZOYER; ROUDART, 2010).

Entende-se por leite bovino, sem outra especificação, o produto oriundo da ordenha completa, ininterrupta, em condições de higiene, de vacas sadias, bem alimentadas e descansadas (BRASIL, 2011).

A exploração da pecuária leiteira pode ser desenvolvida de várias formas: de maneira rústica e em pequenas propriedades ou em grandes empreendimentos com grandes quantidades de vacas e uso de alta tecnologia (SEBRAE, 2015).

O agronegócio leiteiro em Rondônia envolve mais de 38 mil propriedades rurais, emprega 100 mil pessoas nas propriedades rurais e 5 mil na indústria. Apresenta uma produção diária de 2,6 milhões de litros de leite, colocando o estado em 8º lugar na produção de leite nacional e em primeiro da região Norte, sendo uma atividade de relevância social e econômica (SEBRAE, 2015).

A produção de leite é uma das mais importantes atividades do setor agroindustrial rondoniense. A criação de gado para leite é uma atividade fundamental para o estado e que proporciona renda e suporte para a maioria dos produtores (SEBRAE, 2015).

A análise de séries temporais é uma técnica estatística que tem como objetivo: investigar uma série temporal para realizar previsões de valores futuros da série (MORETTIN & TOLOI, 2004).

Segundo Morettin & Toloí (2004) a análise de série temporal é uma metodologia constantemente utilizada, com o propósito de estabelecer projeções futuras, fundamentadas em teorias de que o comportamento da série, ao longo prazo, consiste em repetições de tendências, sazonalidades e variações aleatórias, similares com ocorrido no passado.

Barbosa et al. (2014), aplicou a metodologia Box & Jenkins com o propósito de obter estimativas de produção de leite para os primeiros meses do ano, o modelo SARIMA proposto obteve bons resultados, apresentando ser uma técnica útil para prever as incertezas e imprevistos associados ao mercado lácteo e também para questões práticas administrativas.

Assim, o presente trabalho tem como objetivo aplicar a análise de séries temporais, em dados de produção de leite do estado de Rondônia, no período de janeiro de 2009 a dezembro de 2015. Gerando resultados que possam auxiliar os governantes e produtores rurais nas tomadas de decisões e realizar planejamentos estratégicos futuros.

2. REFERENCIAL TEÓRICO

2.1 AGRONEGÓCIOS LEITEIROS

No Brasil o agronegócio do leite e seus derivados exercem um papel de grande prestígio no fornecimento de alimentos e na criação de empregos e renda para a população (VILELA, 2001).

Vilela (2001) afirma que, para o crescimento da produtividade do leite, serão necessários os investimentos em tecnologias, no melhoramento genético dos rebanhos leiteiros, na alimentação e na saúde animal.

No processo de inovação em tecnologias, por exemplo, a utilização de tratores para preparação do solo e na formação de pastagens, utilização de ordenha mecânica e de tanques de resfriamentos. Essas tecnologias contribuirão para o melhoramento da produtividade do trabalho, alcançando o objetivo principal que é a venda de leite de boa qualidade (DOS SANTOS et al., 2012).

Na região norte, Rondônia concentra a maior parcela dos estabelecimentos rurais que adotam tecnologias de inseminação artificial, ordenha mecanizada e tanques de resfriamento, o que representa um diferencial tecnológico na pecuária leiteira (DOS SANTOS et al., 2012).

2.2 SÉRIES TEMPORAIS

Segundo Morettin & Toloi (2004, pag. 21), os modelos utilizados para descrever séries temporais são processos estocásticos, ou seja, processos controlados por leis probabilísticas.

Uma série temporal é uma coleção de observações feitas sequencialmente ao longo do tempo. A série temporal é dita ser contínua quando as observações são feitas continuamente no tempo. Por outro lado, a série temporal é dita ser discreta quando as observações são feitas em tempos específicos (MORETTIN & TOLOI, 2004).

Segundo Morettin & Toloi (2004), o objetivo da análise de séries temporais é investigar os mecanismos geradores da série temporal, fazer previsões, descrever os comportamento da série e verificar possíveis periodicidades relevantes nos dados em estudo. A metodologia proposta por Box & Jenkins (1994), será utilizada para tais estudos.

A Figura 1 resume a abordagem iterativa para modelo de construção para a previsão proposta por Box & Jenkins (1994).

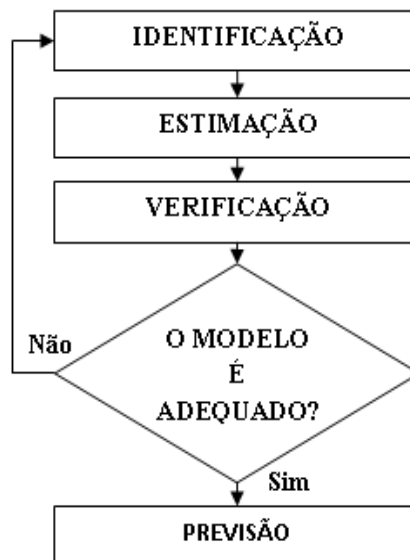


Figura 1. Ciclo iterativo de Box & Jenkins

Fonte: Adaptado de Gujarati & Porter (2011, p. 772).

De acordo Morettin & Toloi (2004) e Ehlers (2007) o procedimento de identificação consiste nas seguintes etapas:

- Verificar se a série necessita de transformações, com o objetivo de estabilizar a variância;
- Tomar as diferenças necessárias até torná-la uma série estacionária. O número de diferenças, d , necessárias para que os dados tornem estacionário;
- Seguidamente estimar os valores de p , número de termos autoregressivos, P , número de termos autorregressivos sazonal, q , número de termos média móvel e

Q, número de termos média móvel sazonal, devem ser especificados com base nas funções de autocorrelação (ACF) e autocorrelação parcial (PACF) da série diferenciada.

Identificando a presença de sazonalidade na série temporal, opta-se por utilizar o modelo sazonal multiplicativo ARIMA (*multiplicative seasonal autoregressive integrated moving average model*), também conhecido como SARIMA (*seasonal autoregressive integrated moving average*), de ordem (p,d,q) x (P,D,Q)s (BOX & JENKINS, 1994).

$$\Phi_P(B^S)\phi_p(B)\nabla_S^D\nabla^d Z_t = \theta_q(B)\Theta_Q(B^{12})a_t \quad (2.1)$$

onde,

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p), \quad (2.2)$$

é o componente autoregressivo de ordem p;

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q), \quad (2.3)$$

é o componente médias móveis de ordem q;

$$\Phi_P(B^S) = (1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_P B^{PS}), \quad (2.4)$$

é o componente sazonal autoregressivo de ordem P;

$$\Theta_Q(B^S) = (1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS}), \quad (2.5)$$

é o componente sazonal médias móveis de ordem Q;

$$\nabla^d = (1 - B)^d, \quad (2.6)$$

é a diferença ordinária de ordem d;

$$\nabla_S^D = (1 - B^S)^D, \quad (2.7)$$

é a diferença sazonal de ordem D;

a_t é ruído branco;

Z_t é a presente observação da série.

Tendo identificado um modelo, o próximo passo é estimar os parâmetros identificados no modelo. Isto é feito por um processo iterativo. Um conjunto inicial de valores para os parâmetros é assumido. Iterativamente um conjunto ideal é obtido (MORETTIN & TOLOI, 2004).

Verificação de diagnóstico, que consiste em, verificar se os resíduos estimados se comportam com ruído branco, ou seja, uma sequência de erros aleatórios de média zero e variância constante onde seguem uma distribuição normal $a_t \sim N(0, \sigma_a^2)$, é realizada através da verificação da autocorrelação e autocorrelação parcial dos coeficientes dos resíduos do modelo (MORETTIN & TOLOI, 2004).

Aplicação desta metodologia poderá gerar um modelo ARIMA para descrever a produção de leite e previsão, em curto prazo, mesmo se a série em estudo apresenta grande variabilidade (SÁNCHEZ-LÓPEZ et al., 2013).

Segundo Souza et al. (2013) os resultados apontaram que o objetivo de prever os valores futuros do leite foi alcançado por meio da metodologia proposta, sendo de grande importância para as empresas que coletam os leites fornecidos pelos pequenos produtores como meio de fornecer informações quanto ao gerenciamento da produção nos períodos de menor e maior produção de leite.

Conforme Sánchez-López et al. (2013) a aplicação desta metodologia também pode ser utilizada pelas organizações de produtores como uma ferramenta adicional para melhor compreender o comportamento de produção de agregados e estimar o seu comportamento futuro, de modo a ter mais informações para apoiar a sua pressão e o processo de planejamento como um grupo criado para apoiar o avanço de sua atividade produtiva.

Gerar previsões para agências governamentais, principalmente em períodos sazonais em que a produção tem variações fortes podem tornar-se um instrumento a ser conhecido pelos produtores e a possibilidade de tomar medidas técnicas para enfrentar ou aproveitar dessas circunstâncias melhorando assim a sua capacidade de resposta.

3. MATERIAIS E MÉTODOS

3.1 BASE DE ESTUDO

Para o desenvolvimento desse estudo foram utilizadas séries históricas mensais da produção de leite cru do estado de Rondônia, os dados utilizados foram obtidos junto ao

Instituto Brasileiro de Geografia e Estatística (IBGE) através do Sistema IBGE de Recuperação Automática (SIDRA).

A base de dados corresponde a 84 observações mensais, entre os períodos de janeiro de 2009 até dezembro 2015, constituindo uma série temporal do tipo discreta. Para a identificação, estimação e verificação do modelo utilizou-se 78 observações, compreendidas no período de janeiro de 2009 até dezembro de 2014. Os doze meses de 2015, compreendidas entre os meses de janeiro de 2015 até dezembro de 2015, foram reservados para comparação com as previsões e validação do modelo ajustado.

3.2 METODOLOGIA BOX & JENKINS

A análise de séries temporais tem como objetivo principal a realização de previsão. Box & Jenkins estabeleceram uma metodologia que permite que valores futuros de uma série sejam previstos tomando por base apenas seus valores do passado e atuais. Isso é feito por meio da correlação temporal existente entre as observações (BOX & JENKINS, 1994).

A primeira etapa para análise de uma série temporal é a construção de seu gráfico, onde poderá ser identificadas características importantes, como por exemplo, tendência, sazonalidade, variabilidade, observações incomuns, ou seja, valores discrepantes (“outliers”) etc (MORETTIN & TOLOI, 2004).

Para a escolha dos modelos seguindo a metodologia proposta é necessário que a série seja estacionária, para que isso ocorra será necessário transformar os dados originais, ou seja, tomar diferenças sucessivas da série original, até torná-la uma série estacionária, em geral uma ou duas diferenças são suficientes (MORETTIN & TOLOI, 2004).

Os modelos ajustados pela metodologia Box & Jenkins (1994) são denominados ARIMA modelo sazonal multiplicativo de ordem $(p,d,q) \times (P,D,Q)_s$, em que d e D são o número de diferenciações necessárias para tornarem a série estacionárias e sem sazonalidade, respectivamente, no geral uma ou duas diferenciações, p é o número de termos autorregressivos, q é o número de termos de médias móveis, P e Q são os parâmetros autorregressivos e média móvel sazonais, s é o período sazonal (BOX & JENKINS, 1994; MORETTIN & TOLOI, 2014).

Para analisar a presença de tendência na série, foi utilizado o teste não paramétrico Cox-Stuart (teste do sinal), nesse teste as observações são agrupados em pares $(Z_1, Z_{1+c}), (Z_2, Z_{2+c}), \dots, (Z_{N-c}, Z_N)$ (MORETTIN & TOLOI, 2004):

onde,

$$(Z_{N-c}, Z_N) = \begin{cases} C = \frac{N}{2}, & \text{se } N \text{ for par} \\ C = \frac{N+1}{2}, & \text{se } N \text{ for impar} \end{cases} \quad (3.1)$$

A cada par (Z_i, Z_{i+c}) é associado um sinal (+ ou -), eliminando os empates. Sendo n o número de pares onde $(Z_i \neq Z_{i+c})$:

$$(Z_i, Z_{i+c}) = \begin{cases} +, & \text{se } Z_i < Z_{i+c} \\ -, & \text{se } Z_i > Z_{i+c} \end{cases} \quad (3.2)$$

Hipóteses testadas:

$H_0 : P(Z_i < Z_{1+c}) = P(Z_i > Z_{1+c}), \forall i$: não existe tendência.

$H_1 : P(Z_i < Z_{1+c}) \neq P(Z_i > Z_{1+c}), \forall i$: existe tendência.

Admitindo T_2 como sendo o número de pares com o sinal +, para $n \leq 20$, a regra de decisão é baseada na distribuição binomial e para $n > 20$, utiliza a aproximação normal.

A identificação de sazonalidade da série foi verificada por meio do teste de Fisher (SILVA et al., 2011; BARBOSA et al., 2014).

Foram testadas as seguintes hipóteses:

H_0 : Não existe sazonalidade.

H_1 : Existe sazonalidade.

Estatística de teste é dada pela seguinte equação:

$$g = \frac{\max I_j^{(N)}}{\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(N)}} \quad (3.3)$$

onde, I é o valor do periodograma no período j e N é o numero de observações da séries. A distribuição exata de g , é dada por $Z_\alpha = 1 - \left(\frac{\alpha}{2}\right)^{\frac{1}{c-1}}$ sendo $c = \frac{N}{2}$ e α o nível de significância. Se $g > Z$, a hipótese H_0 é rejeitada confirmando a presença de sazonalidade (BARBOSA et al., 2014; TIDRE et al., 2013).

A estacionariedade da série será analisada pelo teste Dickey e Fuller Aumentado (ADF), que envolve a estimação pelo método de Mínimos Quadrados Ordinários (MQO). A hipótese nula (H_0), de que a série testada possui raiz unitária, ou seja, não estacionária (CRYER & KELLET, 1986).

Os procedimentos utilizados para a obtenção do modelo proposto são embasados na metodologia proposta por Box & Jenkins (1994) seguem um ciclo iterativo composto pelas seguintes etapas: identificação, estimação, diagnóstico dos resíduos do modelo (verificação) e previsão (BOX & JENKINS, 1994; SOUZA et al., 2013).

O diagnóstico da série compreende em verificar se os resíduos estimados seriam um ruído branco. Para isso, utilizou os testes estatísticos Box-Pierce e Kolmogorov-Smirnov (BOX & PIERCE, 1970; DE MESQUITA LOPES et al., 2013)

3.3 CRITÉRIOS PARA SELEÇÃO E DESEMPENHO DOS MODELOS

Como critério para escolha do melhor modelo utilizou as estatísticas: critério de informação de Akaike (AIC) e erro absoluto médio percentual (MAPE).

A identificação do modelo foi obtida através do Critério de Informação de Akaike (AIC), segundo Flores (2009) e Ehlers (2007) é uns dos modelos mais utilizados em séries temporais. Os cálculos são representados pela equação (3.4).

$$AIC = -2\log \text{verossimilhança maximizada} + 2m \quad (3.4)$$

onde, m é o número de parâmetros ajustáveis do modelo proposto (EHLERS, 2007).

O erro absoluto médio percentual (MAPE) é utilizado como critério de comparação para modelos sugeridos de acordo com os dados em estudo, baseando nos valores reais e estimados (SÁNCHEZ-LÓPEZ et al., 2013; BARBOSA et al., 2014).

3.4 SOFTWARE R

Para as análises estatísticas dos dados, será utilizado o software R.3.2.5 (R DEVELOPMENT CORE TEAM, 2015) através de procedimentos disponíveis em pacotes aplicados à análise de séries temporais. Para as análises dos testes estatísticos, será considerado o nível de significância usual de 5% ($\alpha=0,05$). Os scripts dos procedimentos constam no Apêndice A.

4. RESULTADOS E DISCUSSÕES

Na Figura 2 é apresentado o gráfico da série original das produções mensais de leite cru de Rondônia, em litros, no período de janeiro 2009 até dezembro de 2014.

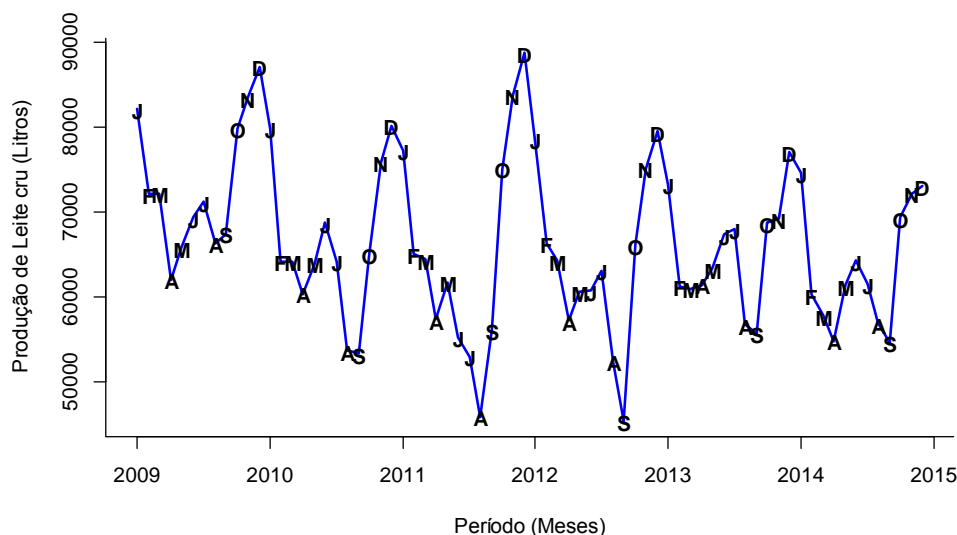


Figura 2. Representação gráfica da série original da produção de leite cru do estado de Rondônia, em litros, no período de janeiro 2009 até dezembro de 2014, com os respectivos meses identificados.

Visualmente (Figura 2), não é possível identificar a presença de tendência, porém é nítido a presença de um comportamento sazonal, com períodos de queda, durante os meses de julho a setembro, e aumento durante os meses de outubro a janeiro. Souza et al. (2010) explicam que no verão, devido à intensidade de chuvas existe abundância da pastagem. Isso permite que o rebanho consumisse maiores quantidades de pasto, consequentemente a um aumento na produção de leite.

Os correlogramas das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) são apresentados na Figura 3, e indicam a presença de sazonalidade.

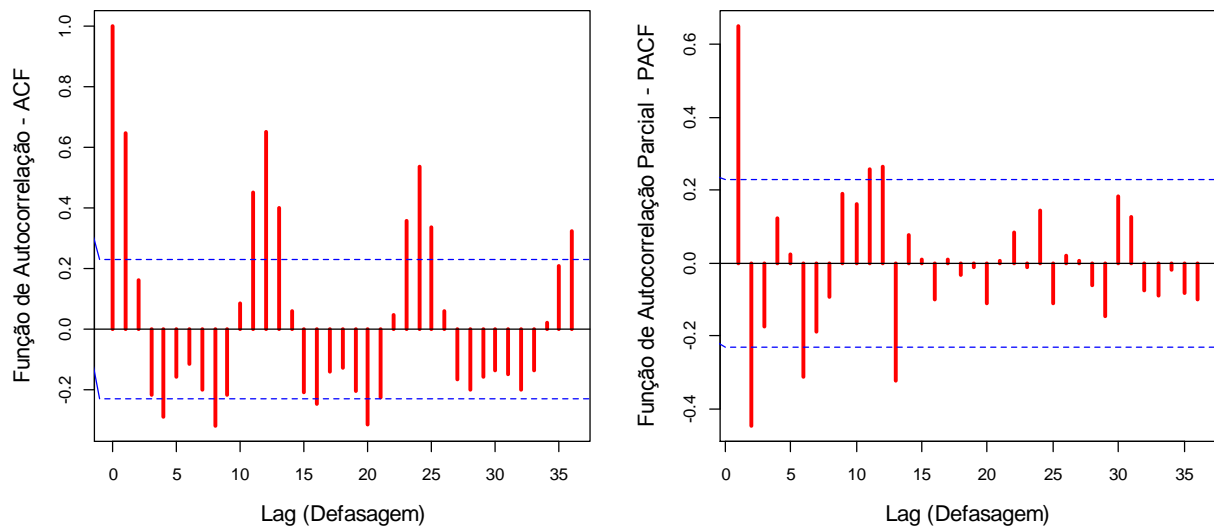


Figura 3. Representação gráfica das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) da série original da produção de leite cru do estado de Rondônia, em litros, no período de janeiro 2009 até dezembro de 2014.

Para confirmar a não existência do componente de tendência foi aplicado o teste Cox-Stuart. A hipótese de nula para a não existência de tendência, não foi rejeitada ao nível de significância igual a 5% ($p\text{-valor} = 0.9994$). A presença do comportamento de sazonalidade foi confirmada pelo teste de Fisher, onde foi testando a hipótese nula para a não presença de sazonalidade, foi rejeitada ao nível de significância igual a 5% ($p\text{-valor} = 0.006$).

Após análises preliminares realizadas, confirmou-se que a série, possui apenas sazonalidade. Logo, torna-se necessário realizar apenas uma diferença na defasagem sazonal da série, a fim de torna-la sem a sazonalidade.

Para a confirmação da estacionariedade antes e após a diferença sazonal na defasagem 12, aplicou-se o teste de Dickey e Fuller Aumentado (ADF), a hipótese nula de raiz unitária foi rejeitada a um nível significância igual a 5% ($p\text{-valor} = 0.03025$).

Segue abaixo os correlogramas da série da produção de leite cru do estado de Rondônia, Figuras 4.

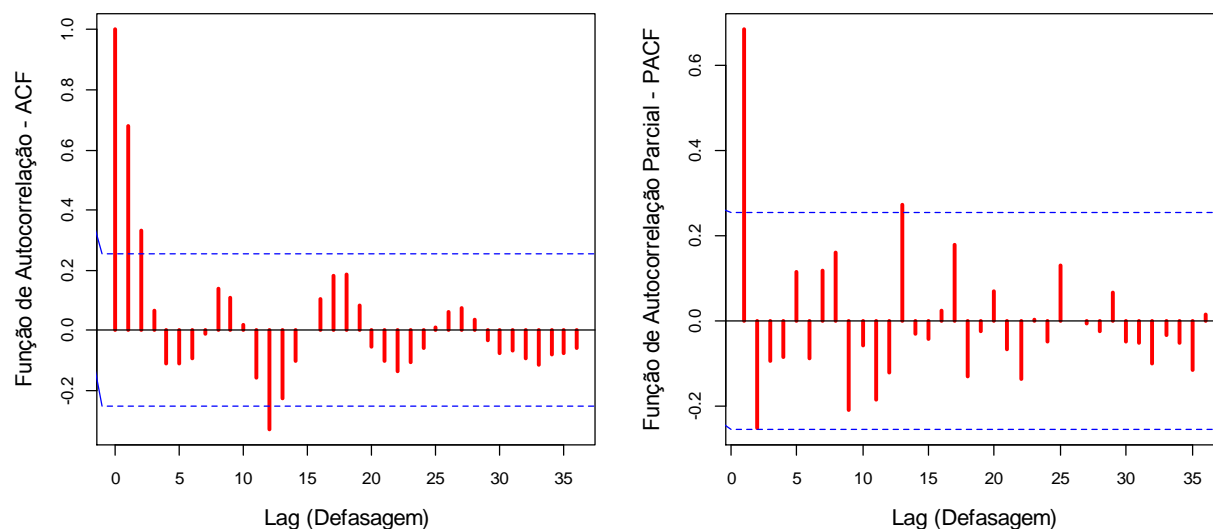


Figura 4. Representação gráfica das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) da série diferenciada sazonalmente da produção de leite cru de Rondônia, em litros, no período de janeiro 2009 até dezembro de 2014.

Após análise visual dos correlogramas da Figura 4, aplicou-se a metodologia proposta por Box & Jenkins, ajustando modelos SARIMA (p,d,q) x (P, D, Q)_s. Foi aplicada uma diferença de ordem 12. Os modelos sugeridos foram:

- SARIMA (1,0,0) x (1,1,0)₁₂
- SARIMA (1,0,0) x (0,1,0)₁₂

Na Tabela 1, segue os testes dos resíduos para independência e normalidade, verificados dos modelos propostos.

Tabela 1. Resultados dos *p*-valor para os testes dos resíduos.

Teste dos Resíduos			
Modelos	SARIMA(p, d, q)x(P, D, Q) ₁₂	<i>p</i> -valor	
		Box-Pierce*	Kolmogorov-Smirnoff
1	(1,0,0)x(1,1,0) ₁₂	0.5145	0.1089
2	(1,0,0)x(0,1,0) ₁₂	0.0510	0.2592
* Calculado para até as primeiras 24 autocorrelações			

O teste de Box-Pierce mostrou que os resíduos são independentes e identicamente distribuídos, com média zero e variância constante, para os dois modelos sugeridos.

O teste de Kolmogorov-Smirnoff confirmou a existência de normalidade dos resíduos, ao nível de significância de 5%.

Na Tabela 2, seguem as estimativas dos modelos sugeridos, identificação do melhor modelo.

Tabela 2. Estatísticas dos modelos SARIMA (p,d,q) x (P,D,Q)s.

SARIMA (1, 0, 0) x (1, 1, 0)₁₂				
Parâmetro	Coefficiente	Erro Padrão	Intervalo de confiança	
			2,5%	97,5%
AR1	0.7746	0.0852	0.5776	0.9115
SAR1	-0.4491	0.1154	-0.6753	-0.2229
AIC = 1175.6; MAPE= 4.075573				
SARIMA (1, 0, 0) x (0, 1, 0)₁₂				
Parâmetro	Coefficiente	Erro Padrão	Intervalo de confiança	
			2,5%	97,5%
AR1	0.7025	0.089	0.5280	0.8769
AIC = 2335.44; MAPE= 4.61415				

De acordo com valores na Tabela 2, verificou-se que o modelo SARIMA (1,0,0) x (1,1,0)₁₂, destacou-se como o melhor modelo ou mais adequado, pois apresentou os menores valores para AIC e MAPE. Analisando o intervalo de confiança em nenhum momento os parâmetros do modelo assumiram o valor zero.

Na Figura 5 tem-se uma representação gráfica das previsões encontradas pelos dois modelos SARIMA.

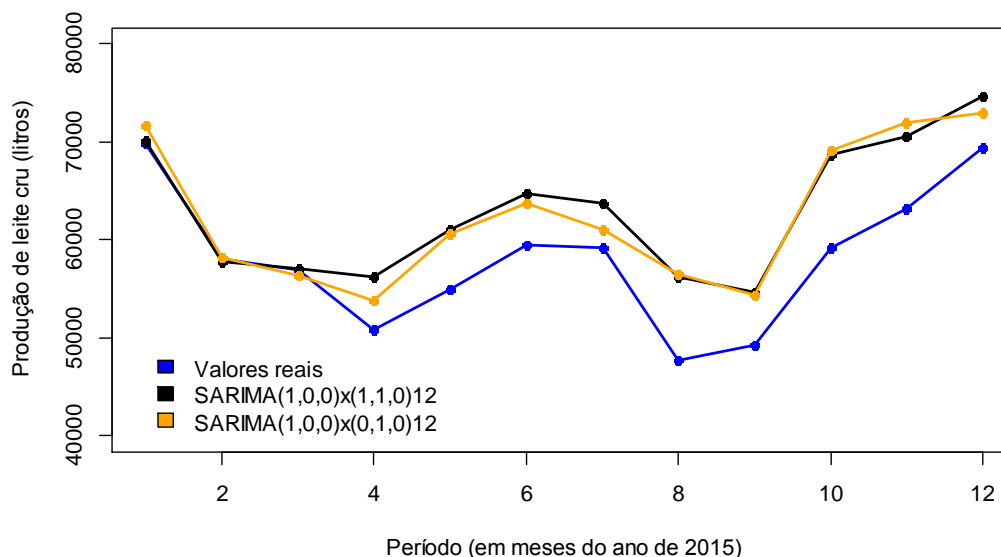


Figura 5. Comparação entre os valores reais e previstos da série da produção de leite cru para o estado de Rondônia para o período de janeiro de 2015 até dezembro de 2015.

Visualmente os dois modelos propostos, apresentaram uma boa previsão em curto prazo (Figura 5 e Tabela 3), até o mês de março de 2015, a partir de abril os dois modelos previstos estão superestimando os valores reais, ou seja, estão sempre acima, porém o modelo escolhido SARIMA (1,0,0) x (1,1,0)₁₂ obteve os melhores resultados.

Tabela 3. Valores reais e valores previstos do modelo SARIMA (1,0,0) x (1,1,0)₁₂, no período de janeiro a dezembro de 2015, da produção mensal de leite de Rondônia (em litros).

Período	Limite inferior 5%	Produção prevista (litros)	Limite superior 95%	Produção real (litros)	Erro de previsão (litros)
jan/15	62278.01	70154.71	78031.4	69803	351.71
fev/15	48025.66	57845.99	67666.32	58002	-156.01
mar/15	46333.23	57080.72	67828.21	56970	110.72
abr/15	44965.9	56194.47	67423.04	50890	5304.47
mai/15	49571.23	61057.83	72544.44	54982	6075.83
jun/15	53120.56	64747.75	76374.94	59421	5326.75
jul/15	51972.83	63677.24	75381.64	59227	4450.24
ago/15	44526.49	56273.49	68020.48	47658	8615.49
set/15	42945.68	54716.22	66486.76	49232	5484.22
out/15	56973.89	68757.46	80541.03	59204	9553.46
nov/15	58778.85	70569.64	82360.44	63180	7389.64
dez/15	62814.09	74608.89	86403.68	69385	5223.89

Nota-se que os doze valores previstos de produção de leite para o ano de 2015 no estado de Rondônia situam-se dentro do nível de confiança de 95% fornecidos pelo modelo ajustado.

Para os primeiros três meses de previsões de produção de leite o modelo obteve um comportamento adequado, para o mês de jan/15 foi de 70154.71 mil litros de leite, com limite inferior de 62278.01 mil litros e superior de 78031.4 mil litros, para o mês de fev/15 produção total de 57845.99 mil litros com limites inferior e superior entre 48025.66 a 67666.32 mil litros, respectivamente.

De jan/15 a mar/15 foram os meses que apresentaram os menores erros de previsão, comparando com os outros meses previstos, embora todos os meses estimados de 2015 estivessem dentro do nível de confiança.

5. CONCLUSÕES

A metodologia Box & Jenkins foi empregada no objetivo de obter estimativas de produção de leite para os doze primeiros meses de 2015. O modelo SARIMA $(1,0,0) \times (1,1,0)_{12}$ obteve os melhores resultados, apresentado o menor erro de acordo com os indicadores selecionados, além de resíduos com características de ruído branco, o que corrobora a qualidade do ajuste.

Observa-se que os doze valores previstos de produção de leite para o ano de 2015 no estado de Rondônia situam-se dentro do nível de confiança de 95% fornecidos pelo modelo ajustado. Os resultados encontrados são bons, principalmente devido às incertezas e imprevistos associados ao mercado lácteo e também para questões práticas, que envolvem tomadas de decisões e a realização de planejamentos estratégicos futuros.

REFERÊNCIAS

- BARBOSA, E. C.; SÁFADI, T.; SILVA, C. H. O.; MANULI, R. C. Metodologia Box & Jenkins: uma aplicação em dados de produção de leite cru do estado de Minas Gerais. **Revista do Instituto de Laticínios Cândido Tostes**, v. 69, n. 2, p. 129-139, 2014
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. Instrução Normativa nº 62, de 29 de dezembro de 2011. **Regulamento técnico de produção, identidade e qualidade do leite tipo A, regulamento técnico de identidade e qualidade de leite cru refrigerado, regulamento técnico de identidade e qualidade de leite pasteurizado e o regulamento técnico da coleta de leite cru refrigerado e seu transporte a granel**. Diário Oficial da União, 30 dez. 2011. Seção 1.
- BOX, G. E. P., JENKINS, G. M., REINSEL, G. C., (1994) **Time Séries Analysis Forecasting and Control**, Prentice Hall, New Jersey, 3rd edition.
- BOX, G. E. P.; PIERCE, D. Distribution of residual autocorrelations in autoregressive integrated moving average time séries models. **Journal of the American Statistical Association**, v 65, p. 1509-1526, 1970.
- CRYER, J. D.; KELLET, N. **Time séries analysis**. Boston: Duxbury Press, 1986.
- DE MESQUITA LOPES, M.; BRANCO, V. T. F. C.; SOARES, J. B. Utilização dos testes estatísticos de Kolmogorov-Smirnov e Shapiro-Wilk para verificação da normalidade para materiais de pavimentação. **Transportes**, v. 21, n. 1, p. 59-66, 2013.
- DOS SANTOS, M. A. S.; DE SANTANA, A. C.; RAIOL, L. C. B. Índice de modernização da pecuária leiteira no estado de Rondônia: determinantes e hierarquização. **Perspectiva Econômica**, v. 7, n. 2, p. 93-106, 2012.
- EHLERS, Ricardo S. Análise de séries temporais. **Laboratório de Estatística e Geoinformação. Universidade Federal do Paraná**, 2007.
- FLORES, João Henrique Ferreira. Comparação de modelos MLP/RNA e modelos Box-Jenkins em séries temporais não lineares. 2009.
- GUJARATI, D. N.; PORTER, D. C. **Econometria básica**. 5 ed. Porto Alegre: AMGH, 2011. 59
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Sistema IBGE de Recuperação Automática (SIDRA)**: banco de dados agregados. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 03 nov. 2015
- MAZOYER, M.; ROUDART, L. História das agriculturas no mundo: do neolítico à crise contemporânea. [Tradução de FERREIRA, C. F.F. B]. São Paulo: Editora UNESP; Brasília, DF: NEAD, 2010, 568pg: il. Disponível em <<http://w3.ufsm.br/gpet/files/Historia%20das%20agriculturas%20no%20mundo%20-%20Mazoyer%20e%20Roudart.pdf>> Acesso em: 16 de março de 2016.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. São Paulo: Edgard Blucher, 2004.

R DEVELOPMENT CORE TEAM (2015). **R: a language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. Disponível em: < <http://www.r-project.org> >. Acesso em: out de 2015.

SÁNCHEZ-LÓPEZ, E.; BARRERAS-SERRANO, A.; PÉREZ-LINARES, C.; FIGUEROSA-SAAVEDRA, F.; OLIVAS-VALDEZ, J. A. Aplicación de un modelo arima para pronosticar la producción de leche de bovino en Baja California, MÉXICO. **Tropical and Subtropical Agroecosystems**, v. 16, n. 3, p. 315-324, 2013.

SEBRAE, Serviço de Apoio às Micro e Pequenas Empresas em Rondônia. Diagnóstico do Agronegócio do Leite e Derivados do Estado de Rondônia. Porto Velho: **SEBRAE**, 2015.

SILVA, R. B. V.; FERREIRA, D. F.; SÁFADI, T. Modelos de séries temporais aplicados à série dos índices de preços ao consumidor na região de Lavras, MG, no período de 1992 a 1999. **Organizações Rurais & Agroindustriais**, v. 2, n. 2, 2011.

SOUZA, A. M.; LÍRIO, G. W.; KLIDZIO, R. Análise da estabilidade da produção de leite no município de Santo Ângelo - RS. **Ciência e Natura**, v. 34, n. 2, p. 07-26, 2013.

TIDRE, P. V. V.; BIASE, N. G. G.; DE SOUSA SILVA, M. I. Utilização dos modelos de séries temporais na previsão do consumo mensal de energia elétrica da região Norte do Brasil. **Matemática e Estatística em Foco**, v. 1, n. 1, p. 57-66, 2013.

VILELA, Duarte. A importância econômica, social e nutricional do leite. **Revista Batavo**, n. 111, 2001.

APÊNDICE A - ROTINA DA ANÁLISE DE SÉRIES TEMPORAIS

```
## Limpar a memória do R
rm(list=ls()); ls()

## Mudar diretório
setwd("D:/BK Karla/Documentos/Documents/Unir Materias/TCC");getwd()

## Chamar os pacotes
library(tséries) ; library(MASS) ; library(forecast) ; library(timeSéries) ; library(nortest)

## Chamar os dados
leite<-read.table("leite009.txt", header=F) ; dim(leite)

## Abrir uma nova janela dos gráficos
win.graph(width=8,height=5,pointsize=12)

## Colocar os dados em séries temporais
x=ts(leite,frequency=12,start=c(2009,1))
plot(x,xlab="Tempo (em anos)",ylab="Produção de Leite cru (Litros)",col='blue', type='o',
bty="l")
x1=ts(x, start=c(2009, 1), end=c(2014, 12), freq=12)
x1 = window(x, start=c(2009,1), end=c(2014,12))
plot(x1,xlab='Período (Meses)',ylab='Produção de Leite cru
(Litros)',,cex.lab=1.3,lwd=2,col='blue', bty="l")
Month=c('J','F','M','A','M','J','J','A','S','O','N','D');
points(window(x1,start=c(2009,1),end=c(2014,12)),pch=Month,font=2)

plot(decompose(x1), col='blue')

## Correlograma da série original
win.graph(width=10,height=5,pointsize=10); par(mfrow=c(1,2))
a=acf(x1,plot=F,na.action = na.fail,lag.max=36);a$lag; a$lag=a$lag*12;
plot(a,main=NA,col="red",cex.lab=1.3,cex.sub=1.4,lwd=3,xlab="Lag
(Defasagem)",ylab="Função de Autocorrelação - ACF")
p=pacf(x1,plot=F,na.action = na.fail,lag.max=36);p$lag; p$lag=p$lag*12;
```

```
plot(p,main=NA,col="red",cex.lab=1.3,cex.sub=1.4,lwd=3,xlab="Lag
(Defasagem)",ylab="Função de Autocorrelação Parcial - PACF")
```

```
## Estatística descritiva
```

```
boxplot( x1 ~ cycle(x1), xlab = "Meses", ylab = "Produção de Leite (Litros)",)
```

```
hist(x1, xlab = "", ylab="", main="")
```

```
## Teste Cox-Stuart - verificar a tendência da série
```

```
cox.stuart=function(x, trend=c("+", "-")){
```

```
  trend=match.arg(trend)
```

```
  N=length(x)
```

```
  p1=x[1:(N/2)]
```

```
  p2=x[(N/2+1):N]
```

```
  sinais=sign(p2-p1)
```

```
  if(trend=="-") T2=sum(sinais<0) else T2=sum(sinais>0)
```

```
  n=length(sinais[sinais!=0])
```

```
  if(trend=="+") tcrit=qbinom(.95,n,.5) else tcrit=qbinom(.05,n, .5)
```

```
  if(trend=="+") pv=1-pbinom(T2, n,.5) else pv=pbinom(T2, n, .5)
```

```
  if(round(pv, 4)==0) pv="<0.0001" else pv=round(pv, 4)
```

```
  data.frame("N"=N, "n"=n, "estatistica"=T2, "valor.critico"=tcrit, "tendencia"=trend, "p-
  valor"=pv)
```

```
}
```

```
cox.stuart(x1,trend="+")
```

```
##Teste Fisher - verificar a sazonalidade da série
```

```
win.graph(width=8,height=5,pointsize=10)
```

```
pgram=spectrum(x1, fast=FALSE, detrend=FALSE, log="no", plot=T, main="Periodograma
da série")$spec
```

```
n=length(pgram)
```

```
pvalue.g=function(g.obs, n) n*(1-g.obs)^(n-1)
```

```
g.obs=max(pgram)/sum(pgram)
```

```
pvalue.g(g.obs,n)
```

```
## Aplicação da primeira diferença sazonal
```

```
k=diff(x1, lag = 12, differences = 1)
```

```
plot(decompose(k), col='blue')
```

```
## Gráfico da série diferenciada
```

```
win.graph(width=8,height=5,pointsize=12)
```

```
plot(k,type="l",cex.lab=1.3,lwd=2,col="blue",xlab="Período (Meses)", ylab="Produção de  
Leite cru (Litros)")
```

```
##Teste Fisher – Para confirmar a eliminação de sazonalidade
```

```
win.graph(width=8,height=5,pointsize=10)
```

```
pgram=spectrum(k, fast=FALSE, detrend=FALSE, log="no", plot=T, main="Periodograma da  
série")$spec
```

```
n=length(pgram)
```

```
pvalue.g=function(g,obs, n) n*(1-g,obs)^(n-1)
```

```
g,obs=max(pgram)/sum(pgram)
```

```
pvalue.g(g,obs,n)
```

```
## Teste ADF - confirmar a estacionariedade da série e ausência de raiz unitária
```

```
adf.test(k, alternative ="stationary")
```

```
## ACF e PACF da série diferenciada
```

```
win.graph(width=10,height=5,pointsize=10); par(mfrow=c(1,2))
```

```
a=acf(k,plot=F,na.action = na.fail,lag.max=36);a$lag; a$lag=a$lag*12;
```

```
plot(a,main=NA,col="red",cex.lab=1.3,cex.sub=1.4,lwd=3,xlab="Lag  
(Defasagem)",ylab="Função de Autocorrelação - ACF")
```

```
p=pacf(k,plot=F,na.action = na.fail,lag.max=36);p$lag; p$lag=p$lag*12;
```

```
plot(p,main=NA,col="red",cex.lab=1.3,cex.sub=1.4,lwd=3,xlab="Lag  
(Defasagem)",ylab="Função de Autocorrelação Parcial - PACF")
```

```
# # Coeficientes dos modelos testados
```

```
mod1=arima(x1, order = c(1,0,0),seas = list(order = c(1,1,0), period=12),method="ML")
```

```
win.graph(width=8,height=5,pointsize=10); tsdiag(mod1); t(confint(mod1));
```

```
res=mod1$residuals
```

```
shapiro.test(res); summary(mod1)
```

```
mod2=arima(x1, order = c(1,0,0),seas = list(order = c(0,1,0), period=12),method="ML")
```

```
win.graph(width=8,height=5,pointsize=10); tsdiag(mod2); t(confint(mod2));
```

```
res=mod2$residuals
```

```
shapiro.test(res); summary(mod2)
```

```
## Teste de Ljung-Box e Box-Pierce até 24 lags - teste (i.i.d) - ruído branco
```



```

Box.test(mod1$residuals, lag=24, type = "Ljung-Box", fitdf=2)
Box.test(mod2$residuals, lag=24, type = "Ljung-Box", fitdf=2)
Box.test(mod1$residuals, lag=24, type = "Box-Pierce", fitdf=2)
Box.test(mod2$residuals, lag=24, type = "Box-Pierce", fitdf=2)

layout(1:2)

m = acf(mod1$residuals, plot=F); m$lag = m$lag*12; plot(m); m1 = pacf(mod1$residuals,
plot=F); m1$lag = m1$lag*12; plot(m1) ;

m2 = acf(mod2$residuals, plot=F); m2$lag = m2$lag*12; plot(m2); m3 =
pacf(mod2$residuals, plot=F); m3$lag = m3$lag*12; plot(m3)

## Teste kolmogorov smirnorf- distribuição normal
ks.test(mod1$residuals, "pnorm", mean(mod1$residuals), sd(mod1$residuals))
ks.test(mod2$residuals, "pnorm", mean(mod2$residuals), sd(mod2$residuals))

## Gráfico com os dois modelos selecionados para comparação
dadosreservados=read.table("D:/BK Karla/Documentos/Documents/Unir
Materias/TCC/dadosreais.txt", quote="\"")
sériereal=ts(dadosreservados,start=c(2015,1),end=c(2015,12),freq=12)
is.ts(sériereal)
sériereal

a = as.data.frame(predict(mod1,12, interval="prediction")); p1=(a$pred)
b = as.data.frame(predict(mod2,12, interval="prediction")); p2=(b$pred)
mes=c(1:12); observado=c(sériereal); e1=c(p1); e2=c(p2)
win.graph(width=8,height=5,pointsize=12)
plot(mes,observado,cex.lab=1,pch=16,type="o",col="blue",lwd=2,xlab="Período (em meses
do ano de 2015)",ylab="Produção de leite cru (litros)",ylim=c(40000,80000))
lines(mes,e1,col="black",lwd=2,type="o",pch=16);
lines(mes,e2,col="orange",lwd=2,type="o",pch=16)
legend(80000,2,c("1o.trim","2o.trim"),col =c("blue","black","orange"), pch=rep(20,2))
legend(locator(1), xpd=TRUE, ncol=1, legend=c("Valores reais",
"SARIMA(1,0,0)x(1,1,0)","SARIMA(1,0,0)x(0,1,0)"),
fill=c("blue","black","orange"),bty="n")

```